

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2001, 18:12:04 ; Search time 18.37 Seconds

(without alignments)
1156.146 Million cell updates/sec

Title: US-09-405-504A-49

Perfect score: 3271

Sequence: 1 MSAITVTLGLFLPLVLA.....MYVPMEDLYNAISAKTLK 620

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt.39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3271	100.0	620	1	VLCG_HUMAN
2	2801	85.6	620	1	VLCG_HUMAN
3	2798	85.5	620	1	VLCG_HUMAN
4	1044	31.9	646	1	FATP_MOUSE
5	1042.5	31.9	646	1	FATP_MOUSE
6	744.5	22.8	623	1	FATL_YEAST
7	398.5	12.2	522	1	CAIC_ECOLI
8	276.5	8.3	561	1	LCFA_ECOLI
9	274.5	8.4	553	1	LCFA_ECOLI
10	272.5	8.3	556	1	LCFA_ECOLI
11	270.5	8.3	563	1	LCFA_ECOLI
12	266.5	8.1	3587	1	LCFA_ECOLI
13	265.5	8.1	544	1	LCFA_ECOLI
14	264.5	8.1	562	1	LCFA_ECOLI
15	261.5	8.0	537	1	LCFA_ECOLI
16	261.5	8.0	544	1	LCFA_ECOLI
17	260.5	8.0	547	1	LCFA_ECOLI
18	259.5	7.9	1274	1	LCFA_ECOLI
19	258.5	7.9	1274	1	LCFA_ECOLI
20	257.5	7.9	561	1	LCFA_ECOLI
21	255.5	7.8	542	1	LCFA_ECOLI
22	255.5	7.8	545	1	LCFA_ECOLI
23	254.5	7.8	545	1	LCFA_ECOLI
24	251.5	7.7	6359	1	LCFA_ECOLI
25	250	7.6	2555	1	LCFA_ECOLI
26	247.5	7.5	4451	1	LCFA_ECOLI
27	246.5	7.5	543	1	LCFA_ECOLI
28	243	7.4	561	1	LCFA_ECOLI
29	238.5	7.3	670	1	LCFA_ECOLI
30	238.5	7.3	6486	1	LCFA_ECOLI
31	237.5	7.3	5255	1	LCFA_ECOLI
32	236	7.2	3587	1	LCFA_ECOLI
33	235.5	7.2	572	1	LCFA_ECOLI

34	233	7.1	560	1	LCFA_ECOLI	P94547 bacillus su
35	229	7.0	550	1	LCFA_ECOLI	P08659 photinus py
36	220	6.7	2607	1	LCFA_ECOLI	P08007 bacillus ii
37	212	6.5	669	1	LCFA_ECOLI	P36333 penicillium
38	212	6.5	3587	1	LCFA_ECOLI	P36333 penicillium
39	211	6.5	653	1	LCFA_ECOLI	P36333 penicillium
40	210.5	6.4	672	1	LCFA_ECOLI	P36333 penicillium
41	210	6.4	662	1	LCFA_ECOLI	P36333 penicillium
42	208.5	6.4	675	1	LCFA_ECOLI	P36333 penicillium
43	207	6.3	661	1	LCFA_ECOLI	P36333 penicillium
44	204.5	6.3	713	1	LCFA_ECOLI	P36333 penicillium
45	203.5	6.2	503	1	LCFA_ECOLI	P36333 penicillium

ALIGNMENTS

RESULT	1	STANDARD	PRT	620 AA.
VLCG_HUMAN				
AC	014975;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	VERY-LONG-CHAIN ACYL-COA SYNTHETASE (EC 6.2.1.-) (VERY-LONG-CHAIN- FATTY-ACYL-COA LIGASE).			
GN	FAVULI OR VLCAS.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RA	Uchiyama A., Aoyama T., Kamijo K., Wakui K., Fukushima Y., Shimozawa N., Suzuki Y., Kondo N., Oritani T., Hashimoto T., Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99216327; PubMed=10198260;			
RA	Steinberg S.J., Wang S.J., Kim D.G., Mihalik S.J., Watkins P.A., Human very-long-chain acyl-CoA synthetase: cloning, topography, and relevance to branched-chain fatty acid metabolism.			
RT	Biochem. Biophys. Res. Commun. 257:615-621(1999).			
RL	1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. PEROXISOMAL. 2- TISSUE SPECIFICITY: EXPRESSED IN LIVER, KIDNEY, PLACENTA AND PANCREAS.			
CC	1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT COVALENT BINDING OF AMP TO THEIR SUBSTRATE.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sdb.ch/announce/ or send an email to license@sdb.ch).			
CC	EMBL: D88308; BAA23644.1; -			
DR	EMBL: AF096290; AAC64973.1; -			
DR	MIM: 603247; -			
DR	InterPro: IPR000873; -			
DR	Pfam: PF00501; AMP-binding; 1.			
DR	PROSITE: PS00455; AMP-BINDING; 1.			
KW	Ligase; Fatty acid metabolism; Peroxisome; Transmembrane.			
FT	TRANSMEM	1	21	POTENTIAL.
FT	TRANSMEM	107	127	POTENTIAL.
FT	TRANSMEM	262	282	POTENTIAL.
SO	SEQUENCE	620 AA;	70312 MW;	BFD333ELB67B3DF8 CAC64;

Query Match 100.0%; Score 3271; DB 1; Length 620;
Best Local Similarity 100.0%; Pred. No. 1e-239;

Matches 620; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSAITVYLAGLFLPLVNLCCPFYFODIGFLKYAAVGRVRSYGGRRPRTILRAFL 60
 DB 1 MSAITVYLAGLFLPLVNLCCPFYFODIGFLKYAAVGRVRSYGGRRPRTILRAFL 60
 QY 61 EKAROTPHKPFLLFDELTLYAQRVRSNOVARALHDHGLRGDCVALMGNEPAYWL 120
 DB 61 EKAROTPHKPFLLFDELTLYAQRVRSNOVARALHDHGLRGDCVALMGNEPAYWL 120
 QY 121 WGLVKGACMACLVNIRAKSLHFCOCCGAKVLLVSPLOAAVEELLPSLKDDVSIY 180
 DB 121 WGLVKGACMACLVNIRAKSLHFCOCCGAKVLLVSPLOAAVEELLPSLKDDVSIY 180
 QY 181 YVSRSTNDGIDSLDKVDVSTEPISPSWSEVTFSTPALITYSGTGLPKAMITHQ 240
 DB 181 YVSRSTNDGIDSLDKVDVSTEPISPSWSEVTFSTPALITYSGTGLPKAMITHQ 240
 QY 241 RIMVGTGLTFVSGLRADVDIYITLPEYHSAALLIGHCIYAGATLALRTKFSASQFMD 300
 DB 241 RIMVGTGLTFVSGLRADVDIYITLPEYHSAALLIGHCIYAGATLALRTKFSASQFMD 300
 QY 301 CRKYNTVYQYIGELLRYLCNSPQKPNDRDHVRLALGNGLRGVWRQFVRFEDICYE 360
 DB 301 CRKYNTVYQYIGELLRYLCNSPQKPNDRDHVRLALGNGLRGVWRQFVRFEDICYE 360
 QY 361 FFAATEGNIGFMNARKVAGVAVNYLQKITYDLIKYDEKDEPVDENGYCVRPVKG 420
 DB 361 FFAATEGNIGFMNARKVAGVAVNYLQKITYDLIKYDEKDEPVDENGYCVRPVKG 420
 QY 421 EYGLLVCKITQULPENGVAAGAKOTEKKLRDYFKKGLDYNSGDLMAVDHENTFYFHDR 480
 DB 421 EYGLLVCKITQULPENGVAAGAKOTEKKLRDYFKKGLDYNSGDLMAVDHENTFYFHDR 480
 QY 481 VGDTEFMKGENAVATTEVADTVGLVDFVQEVNNGVAVHDEHRIQMASTIKKENHEFPDG 540
 DB 481 VGDTEFMKGENAVATTEVADTVGLVDFVQEVNNGVAVHDEHRIQMASTIKKENHEFPDG 540
 QY 541 KLFQHIADYLPYSARPRFLRIODTIEITGTFKHKMTLVEEGFNPAVKDALYFLDDTAK 600
 DB 541 KLFQHIADYLPYSARPRFLRIODTIEITGTFKHKMTLVEEGFNPAVKDALYFLDDTAK 600
 QY 601 MYVPMTEDIYNAISAKTKL 620
 DB 601 MYVPMTEDIYNAISAKTKL 620

RESULT 2

VCLS_RAT STANDARD; PRT; 620 AA.

AC P97524;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE VERT-LONG-CHAIN ACYL-COA SYNTHETASE (EC 6.2.1.1-) (VERY-LONG-CHAIN-FATTY-ACID-COA LIGASE).
 GN FACYL OR VLACS.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 RX NCBI_TaxID=10116;
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=WISTAR; TISSUE=Liver;
 RX MEDLINE=97094763; PubMed=8939997;
 RA Uchiyama A., Aoyama T., Kamijo K., Uchida Y., Kondo N., Orl T., Hashimoto T.;
 RT "Molecular cloning of cDNA encoding rat very long-chain acyl-CoA synthetase.";
 CC J. Biol. Chem. 271:30360-30365(1996).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. PEROXISOMAL.
 CC -1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT COVALENT BINDING OF AMP TO THEIR SUBSTRATE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb.ch).

CC EMBL: D85100; BAA12722.1; -
 DR InterPro: IPR000873; -
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP BINDING; 1.
 KW Ligase; Fatty acid metabolism; Peroxisome; Transmembrane.
 FT TRANSMEM 1 21 POTENTIAL.
 FT TRANSMEM 107 127 POTENTIAL.
 FT TRANSMEM 262 282 POTENTIAL.
 SQ SEQUENCE 620 AA; 70693 MW; 6CF9362DC3805526 CRC64;

Query Match 85.6%; Score 2801; DB 1; Length 620;
 Best Local Similarity 82.4%; Pred. No. 3.5e-204;
 Matches 511; Conservative 55; Mismatches 54; Indels 0; Gaps 0;

QY 1 MSAITVYLAGLFLPLVNLCCPFYFODIGFLKYAAVGRVRSYGGRRPRTILRAFL 60
 DB 1 MSAITVYLAGLFLPLVNLCCPFYFODIGFLKYAAVGRVRSYGGRRPRTILRAFL 60
 QY 61 EKAROTPHKPFLLFDELTLYAQRVRSNOVARALHDHGLRGDCVALMGNEPAYWL 120
 DB 61 EKAROTPHKPFLLFDELTLYAQRVRSNOVARALHDHGLRGDCVALMGNEPAYWL 120
 QY 121 WGLVKGACMACLVNIRAKSLHFCOCCGAKVLLVSPLOAAVEELLPSLKDDVSIY 180
 DB 121 WGLVKGACMACLVNIRAKSLHFCOCCGAKVLLVSPLOAAVEELLPSLKDDVSIY 180
 QY 181 YVSRSTNDGIDSLDKVDVSTEPISPSWSEVTFSTPALITYSGTGLPKAMITHQ 240
 DB 181 YVSRSTNDGIDSLDKVDVSTEPISPSWSEVTFSTPALITYSGTGLPKAMITHQ 240
 QY 241 RIMVGTGLTFVSGLRADVDIYITLPEYHSAALLIGHCIYAGATLALRTKFSASQFMD 300
 DB 241 RIMVGTGLTFVSGLRADVDIYITLPEYHSAALLIGHCIYAGATLALRTKFSASQFMD 300
 QY 301 CRKYNTVYQYIGELLRYLCNSPQKPNDRDHVRLALGNGLRGVWRQFVRFEDICYE 360
 DB 301 CRKYNTVYQYIGELLRYLCNSPQKPNDRDHVRLALGNGLRGVWRQFVRFEDICYE 360
 QY 361 FFAATEGNIGFMNARKVAGVAVNYLQKITYDLIKYDEKDEPVDENGYCVRPVKG 420
 DB 361 FFAATEGNIGFMNARKVAGVAVNYLQKITYDLIKYDEKDEPVDENGYCVRPVKG 420
 QY 421 EYGLLVCKITQULPENGVAAGAKOTEKKLRDYFKKGLDYNSGDLMAVDHENTFYFHDR 480
 DB 421 EYGLLVCKITQULPENGVAAGAKOTEKKLRDYFKKGLDYNSGDLMAVDHENTFYFHDR 480
 QY 481 VGDTEFMKGENAVATTEVADTVGLVDFVQEVNNGVAVHDEHRIQMASTIKKENHEFPDG 540
 DB 481 VGDTEFMKGENAVATTEVADTVGLVDFVQEVNNGVAVHDEHRIQMASTIKKENHEFPDG 540
 QY 541 KLFQHIADYLPYSARPRFLRIODTIEITGTFKHKMTLVEEGFNPAVKDALYFLDDTAK 600
 DB 541 KLFQHIADYLPYSARPRFLRIODTIEITGTFKHKMTLVEEGFNPAVKDALYFLDDTAK 600
 QY 601 MYVPMTEDIYNAISAKTKL 620
 DB 601 MYVPMTEDIYNAISAKTKL 620

RESULT 3

VCLS_MOUSE STANDARD; PRT; 620 AA.

AC 035488;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE VERT-LONG-CHAIN ACYL-COA SYNTHETASE (EC 6.2.1.1-) (VERY-LONG-CHAIN-FATTY-ACID-COA LIGASE).
 GN FACYL OR VLACS.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 RX NCBI_TaxID=10116;
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=WISTAR; TISSUE=Liver;
 RX MEDLINE=97094763; PubMed=8939997;
 RA Uchiyama A., Aoyama T., Kamijo K., Uchida Y., Kondo N., Orl T., Hashimoto T.;
 RT "Molecular cloning of cDNA encoding rat very long-chain acyl-CoA synthetase.";
 CC J. Biol. Chem. 271:30360-30365(1996).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. PEROXISOMAL.
 CC -1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT COVALENT BINDING OF AMP TO THEIR SUBSTRATE.

DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE VERY-LONG-CHAIN ACYL-COA SYNTHETASE (EC 6.2.1.-) (VERY-LONG-CHAIN-
 DE FATTY-ACYL-COA LIGASE)
 GN FACL1 OR VLACS OR VLCS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Liver;
 RA Kemp S., Lu J.-F., Smith K.D.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. PEROXISOMAL.
 CC -1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN AMP-DEPENDENT
 CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF030311; AAB87982.1; -
 DR InterPro: IPR000873; -
 DR Pfam: PF00501; AMP-binding: 1;
 DR PROSITE: PS00455; AMP_BINDING: 1;
 DR L1gase; Fatty acid metabolism; Peroxisome; Transmembrane.
 FT TRANSMEM 1 21 POTENTIAL.
 FT TRANSMEM 107 127 POTENTIAL.
 FT TRANSMEM 262 282 POTENTIAL.
 SQ SEQUENCE 620 AA; 70366 MW; 77C9BDD0E3B9FFB CRC64;

Query Match 85.5%; Score 2798; DB 1; Length 620;
 Best Local Similarity 82.6%; Pred. No 5,9e-204;
 Matches 512; Conservative 52; Mismatches 56; Indels 0; Gaps 0;

QY 1 MSAIYVLAGLPLPLVNLCPYFODIGYELKVAAGRRVRSQGRPARTILRAFL 60
 Db 1 MLPVLYGLAGLPLPLPLFCCEYLLQDVYFQLANMARVRVRSQGRPARTILRAFL 60
 QY 61 EKAROTPKPPLRDELRTTAQVDRSRNOYARALHDLGLRGDCAVALLMGNEPAVWL 120
 Db 61 EQAKRTPKPPLRDELRTTAQVDRSRNOYARALHDLGLRGDCAVALLMGNEPAVWI 120
 QY 121 WLGLVKGICAMACLNINIRAKSLHRCQCCGAKVLLVSPETOAIVEILSLKDDVSIV 180
 Db 121 WLGLLKGICCPACLNINIRAKSLHRCQCCGAKVLLVSPDQEAIVEALPLTKDAVSIF 180
 QY 181 YVSRSTNDGIDSLFKVDEVSTPEDESRSEVTESTPALYITTSCTGLPKRAMITHQ 240
 Db 181 YVSRSTNNGVDLIDVDYSAEPTESEMSSEVTEFTPAVYITTSCTGLPKRAMITNH 240
 QY 241 RIWGTGTFVSGIKADVDYITLPEYHSALLIGISVAGTALTALRKFSQSCWMD 300
 Db 241 RLVGTGLAMSSGITADVDYITLPEYHSALLIGISVAGTALTALRKFSQSCWMD 300
 QY 301 CRKYNVTVIOYIGELRLYCLNSPOKPNDRDHKVALGNGLRGVDWRFKREGDLCIYE 360
 Db 301 CRKYNVTVIOYIGELRLYCLNSPOKPNDRDHKVALGNGLRGVDWRFKREGDHIYE 360
 QY 361 FYATTEGNTGMYARKVAGVGRVNYLQKTIITDILKIYVEKDEPVNDGTCVVRPKG 420
 Db 361 FYASTEIGIGVNPRIKIGAVGRVNYLQKRYARVYLKIYVEKDEPVNDGTCVVRPKG 420
 QY 421 EVGLLVCKITQLPFNGYAGAKAOTEKKLRDVEKGDLEFNSGDLMLVHENVYFHDR 480
 Db 421 EVGLLVCKITQLPFNGYAGAKAOTEKKLRDVEKGDLEFNSGDLMLVHENVYFHDR 480

QY 481 VGDTEFRKGNVATTEVADTVGLVDVQEVNYYGVAPHDEGRIGMASIKMEHEDGK 540
 Db 481 VGDTEFRKGNVATTEVADTVGLVDVQEVNYYGVAPHDEGRIGMASIKMEHEDGK 540
 QY 541 KLFQHIADYLPYARPRFLRIODTIEITGFKRKKTLEEGPNPAVITDALYPLDDFAK 600
 Db 541 KLFQHIADYLPYARPRFLRIODTIEITGFKRKKTLEEGPNPAVITDALYPLDDFAK 600
 QY 601 MYVPMTEDIYNATSAKTLLK 620
 Db 601 TEVPMTEINIYNAIDKTLK 620

RESULT 4
 FATP_MOUSE
 ID FATP_MOUSE STANDARD; PRT; 646 AA.
 AC 060714;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 DE LONG-CHAIN FATTY ACID TRANSPORT PROTEIN (FATP).
 GN SLC27A1 OR FATP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SWISS;
 RX MEDLINE=95042740; PubMed=7954810;
 RA Schaffer J.E., Lodish H.F.;
 RT "Expression cloning and characterization of a novel adipocyte long
 RT chain fatty acid transport protein.";
 RL Cell 79:427-436(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98438516; PubMed=9765271;
 RA Hul T.Y., Froehner B.I., Smith A.J., Schaffer J.E., Bernlohr D.A.;
 RT "Characterization of the murine fatty acid transport protein gene and
 RT its insulin response sequence.";
 RL J. Biol. Chem. 273:27420-27429(1998).
 CC -1- FUNCTION: INVOLVED IN TRANSLLOCATION OF LONG-CHAIN FATTY ACIDS
 CC ACROSS THE PLASMA MEMBRANE. MAY PLAY A PIVOTAL ROLE IN REGULATING
 CC AVAILABLE LONG-CHAIN FATTY ACID SUBSTRATES FROM EXOGENOUS SOURCES
 CC IN TISSUES UNDERGOING HIGH LEVELS OF BETA-OXIDATION OR
 CC TRIGLYCERIDE SYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. PLASMA MEMBRANE.
 CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN SKELETAL MUSCLE, HEART
 CC AND FAT. LOWER LEVELS IN BRAIN, KIDNEY, LUNG AND LIVER. NO
 CC EXPRESSION IN SPLEEN OR INTESTINE.
 CC -1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT
 CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U15976; AAC71060.1; -
 DR EMBL: AF023258; AAC69640.1; -
 DR EMBL: AF023256; AAC69640.1; JOINED.
 DR EMBL: AF023257; AAC69640.1; JOINED.
 DR MGD: MGI:1347098; SLC27A1.
 DR InterPro: IPR000873; -
 DR Pfam: PF00501; AMP-binding: 1;
 DR PROSITE: PS00455; AMP_BINDING: 1;
 KW Glycoprotein; Lipid transport; Transmembrane.
 FT TRANSMEM 14 34 POTENTIAL.
 FT TRANSMEM 115 135 POTENTIAL.

FT TRANSMEM 140 160 POTENTIAL.
 FT TRANSMEM 169 189 POTENTIAL.
 FT TRANSMEM 293 313 POTENTIAL.
 FT TRANSMEM 357 357 POTENTIAL.
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 393 393 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 646 AA; 71276 MW; 910B92BA8D985B4C CRC64;

Query Match 31.9%; Score 1044; DB 1; Length 646;
 Best Local Similarity 40.0%; Pred. No. 2.5e-71;
 Matches 247; Conservative 83; Mismatches 245; Indels 42; Gaps 11;

33 FLVAAGVRVRVSYG-----ORPARTLRAFLKAROTPKPLRDETL 79
 43 FLRIYVKTRARDFGLSVLIRKRLRRRRAGDTIFCFQAVARQPRRLALVDASSI 102
 80 --TYAQVRRSNQVARALHDHGLRGDCVALLMGNEPAYWLMGLVKGACMACLYN 137
 103 CMTFADLDYTSNAVAN-LEFQIGFAPGDVAVAFLEGRPEFVGLMGLAKAGVALLNVN 161
 138 IIRAKSLHFGCCGKAVLVSPLOAAVEELPSLKDDVSIYYSRTNTGIDISFLDK 197
 162 LRREPLAFCLGTSAAKALITYGGEMAAVAEVSQIGKS-----LKEGSGDGLPESTILPD 216
 198 VD-----EVSTPEI---PESWRSEVTFSTPALITYSGTGLPKAMITHQRIWYGTG 247
 217 TQLLDMLAEAPTPPLAQAQKGMDRLE-----YITSGTGLPKRAALVHSRYRTAA 271
 248 LTFVVS-GLKADVYITLPFYSALLIGHCIVAGATLALTRKFSASQFMDCRKYNV 306
 272 FGHHSYSMRADVLVLCLEPLYSAGNMGVCVIGLIVLRKKFSASRFWDDCKVNC 331
 307 TVYQYIGELLRYLCSNPKRPNDRDHKVRKALNGSLRGDWQRFVAKRGDICIYEYAAE 366
 332 TVVQYIGELCRYLRLQPRVDEQRHVRVLAVNGELRPAIMEETLOFEPVQIGEEYGA 391
 367 GNIGFMNVARKYGAVGRVNYLQKITYLDIKYDEKDEPVDEMGCYRVKGEVGLV 426
 392 CNGSIANNMGKVGSCGFNSRLITHYPRILVKNEDTMEPLDSELCIQCPGPEGLV 451
 427 CKITQLTP--FNGYAGAKAQTEKKRLRVFKKGLDYFNSGDLLENHENTYFHDVGD 483
 452 GQINODDPLRFEDGYSDSA-TNKKLAHSVFRKGSADYVLCSDVLVDELGYMYFDRSGD 510
 484 TFRWKEGNAVTEPAVDYGLVDEYVNYGVHPDHEGRIGMASIKMKNENEPGKILF 543
 511 TFRWKEGNAVTEPAVDYGLVDEYVNYGVHPDHEGRIGMASIKMKNENEPGKILF 569
 544 OHIADYLPSTARPRFLRDIOTTEITGTFKHKRKTVEEGFNPVAVIKDALYFLDDTAKMY 603
 570 QELQKVLASARPIFLRLRPOVDYTGTFKIQTRLRRESFDRQSDRLFFLDMQGRV 629
 604 PMTEDIYNAISAKTKL 620
 630 PLDERVHARICAGDFSL 646

RESULT 5
 FARP_RAT STANDARD; PRT; 646 AA.

AC P97849;
 DT 01-NOV-1997 (rel. 35, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DT 30-MAY-2000 (rel. 39, Last annotation update)
 DE LONG-CHAIN FATTY ACID TRANSPORT PROTEIN (FARP).
 GN SLG27AI OR FARP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=98041635; PubMed=9375787;
 RA Schap F.G., Hamers L., van der Vusse G.J., Glatz J.F.C.;
 RL "Molecular cloning of fatty acid-transport protein cDNA from rat.";
 RL Biochim. Biophys. Acta 1354:29-34(1997).
 CC -1- FUNCTION: INVOLVED IN TRANSLLOCATION OF LONG-CHAIN FATTY ACIDS
 CC ACROSS THE PLASMA MEMBRANE. MAY PLAY A PIVOTAL ROLE IN REGULATING
 CC AVAILABLE LONG-CHAIN FATTY ACID SUBSTRATES FROM EXOGENOUS SOURCES
 CC IN TISSUES UNDERGOING HIGH LEVELS OF BETA-OXIDATION OR
 CC TRIGLYCERIDE SYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. PLASMA MEMBRANE.
 CC -1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT
 CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: U89529; AAC53424.1; -
 DR InterPro: IPR000873; -
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 KW Glycoprotein; Lipid transport; Transmembrane.
 FT TRANSMEM 14 34 POTENTIAL.
 FT TRANSMEM 115 135 POTENTIAL.
 FT TRANSMEM 140 160 POTENTIAL.
 FT TRANSMEM 169 189 POTENTIAL.
 FT TRANSMEM 293 313 POTENTIAL.
 FT TRANSMEM 357 357 POTENTIAL.
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 393 393 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 646 AA; 71283 MW; C450CF14CC2EB87 CRC64;

Query Match 31.9%; Score 1042.5; DB 1; Length 646;
 Best Local Similarity 39.7%; Pred. No. 3.2e-71;
 Matches 251; Conservative 87; Mismatches 259; Indels 35; Gaps 12;

5 IYTVLAGLFLPLVNLCCPFYFQDYGFLVAAGVRVSYGQRRPARTILRAFLKAR 64
 34 YVVGSGGWRFLRIYCKTARBDLF-GLSVLIRVLELR-----HRRAGDTIPRIFQAVAG 87
 65 QTPKPELRFDETL--TYAQVRRSNQVARALHDHGLRGDCVALLMGNEPAYWLM 122
 88 RQPRRLALVDASSGICMTFQDLDTYSNAVAN-LEFQIGFAPGDVAVAFLEGRPEFGLWL 146
 123 GLVTLGCMACMLNINIRAKSLHFGCCGKAVLVSPLOAAVEELPSLKDDVSIYV 182
 147 GLAKAGVALLNVNLRRREPLAFCLGTSAAKALITYGGEMAAVAEVSQIGKS-----LL 201
 183 SRTSNTGIDISFLDKVD-----EVSTPEI---PESWRSEVTFSTPALITYSGTGLP 232
 202 KFCGSGDLGPESVLPDQQLDMLAEAPTPPLAQAQKGMDRLE-----YITSGTGLP 256
 233 KAAMITHQRIWYGTGLTFVS-GLKADVYITLPFYSALLIGHCIVAGTALTRK 291
 257 KAAIVHSRYRRIARFAGHSYSMRANDVLDCLPLYSAGNMGVCVIGLIVLRKK 316
 292 FSAQFMDCRKYNVYIYQYIGELLRYLCSNPKRPNDRDHKVRKALNGSLRGDWQRFV 351
 317 FSAQFMDCRKYNVYIYQYIGELLRYLCSNPKRPNDRDHKVRKALNGSLRGDWQRFV 376
 352 RFQDICIYEFAATEGNIGFMNVARKYGAVGRVNYLQKITYLDIKYDEKDEPVARDEN 411
 377 GFGVROIGFEFGATGECNSIANNMGKVGSCGFNSRLITHYPRILVKNEDTMEPLRDSQ 436
 412 GYCVRVKGEVGLLVCKITQLTP--FNGYAGAKAQTEKKRLRVFKKGLDYFNSGDL 468

DB 437 GCLIPQGPPEGLLVGOINQDPLRRFDGYSDA-TNKKIAHYERRGDSAYLSGDVLY 495
 QY 469 VDHENIYFHDVGDTEPRKGENNATTEADVGLVDYQVENVYGVAPDHEGHTGAS 528
 DB 496 MDELGMTRFDRSDGTFERRGENSVSTEEVAVLSRLIGOTDVAVGVAVPGEKSGMA 555
 QY 529 IMKENHEFDGKLLFOHIDADYLPSPARPELRIDQTEITGTFKRRKTLVEGEPNPAVI 588
 DB 556 IADPHN-QDDPSMTQGLQKVLASTAQAQPLFLRLLPQVDTGTGFKIQKTRLRQREGFDPQT 614
 QY 589 KALYFLDDTAKMYVPMTEDINYAISAKTLK 620
 DB 615 SDRLEFLDKGGRYLPDERVARICAGDFSL 646

RESULT 6
 PART1_YEAST STANDARD; PRT; 623 AA.
 ID PART1_YEAST STANDARD; PRT; 623 AA.
 AC P38225;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PROBABLE LONG-CHAIN FATTY ACID TRANSPORT PROTEIN.
 GN FATT1 OR YBR041W OR YBR0411.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_Taxid=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C;
 RA Andre B., Cziepluch C., Hein C., Jauniaux J.C., Urrestarazu A.,
 RA Vissers S.;
 RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
 RM [2]
 RP CHARACTERIZATION.
 RC STRAIN=W303A;
 RX MEDLINE=97235810; PubMed=9079682;
 RA Faergeman N.J., Dirusso C.C., Elberger A., Knudsen J., Black P.N.;
 RT "Disruption of the Saccharomyces cerevisiae homologue to the murine
 RT fatty acid transport protein impairs uptake and growth on long-chain
 RT fatty acids." 272:8531-8538(1997).
 CC CC
 CC -1- FUNCTION: MAY BE INVOLVED IN LONG-CHAIN FATTY ACIDS UPTAKE, AND
 CC THIS MAY PLAY A PIVOTAL ROLE IN REGULATING THEIR ACCESSIBILITY
 CC PRIOR TO METABOLIC UTILIZATION. MAY PLAY AN IMPORTANT ROLE IN
 CC UPTAKE OF THESE HYDROPHOBIC COMPOUNDS UNDER CONDITIONS WHERE FATTY
 CC ACID SYNTHESIS IS COMPROMISED.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: TO ENZYMES WHICH ACT VIA AN ATP-DEPENDENT COVALENT
 CC BINDING OF AMP TO THEIR SUBSTRATE.
 CC CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC CC
 CC EMBL: Z35910; CA84983.1; -
 CC PIR: S45899; S45899.
 CC SGD: S0000245; FATT1.
 CC InterPro: IPR000873; -
 CC Pfam: PF00501; AMP-binding; 1.
 CC PROSITE: PS00455; AMP BINDING; 1.
 KW Lipid transport; Transmembrane.
 FT TRANSMEM 6 26 POTENTIAL.
 FT TRANSMEM 54 71 POTENTIAL.
 FT TRANSMEM 149 169 POTENTIAL.
 FT TRANSMEM 293 313 POTENTIAL.
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 534 534 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 623 AA; 71697 MW; 0AE02700B60C8CFE CRC64;
 Query Match 22.8%; Score 744.5; DB 1; Length 623;
 Best Local Similarity 32.8%; Pred. No. 1e-48; Indels 59; Gaps 16;
 Matches 191; Conservative 96; Mismatches 236;
 QY 24 PYEPDIDIGYFLKVAAGR-----RVSYGQRPPARTILRAFLERAKQTPKPF 72
 DB 58 PYFLKSY--FCYILDVARHRRFQNMWYLEFKQVQNGDHL-AISTRRMAEKG----- 106
 QY 73 LFRDETLTYAQNDRSRQVAPALHDHGLGKQSGCVALLMGNEPAYVWLGVLKGCAMA 132
 DB 107 -FOLEPTTYIETNYINYLRLSHILHFDYNNVQAGDYVAIDCTNKLDFVLMLSLNIIGAIPA 165
 QY 133 CLANNIRAKSLHCOFCGAKVLLVSPELQAAVEELPSLKK--DDVSITYYSRTSNTDG 190
 DB 166 FLNNTYKGTPLVSLKLSNITQVFIIDPDASNPIRESEELKNALPDVKLYLEEQ----- 220
 QY 191 IDSLDKVDVSTEEPIPEWSRSE-----VFESTPALYIYSGTGLPKKAMITHQRTM 243
 DB 221 ----DLMHLLNSQSPFELQDQNVKPTPLGLTFKPSMLIYTSGTGLPKSAIMSRKSS 275
 QY 244 YGTGL-TFVSGLRADVDIYITLFFYSALLIGHCYAGATLARTFSAQFDDCR 302
 DB 276 VGCQVGHVLMHTNESTVEFAMPLFHSHTALLGACAILSHGCLASHKFSASTFWKQVY 335
 QY 303 KYNVTYQYIGELRLYLCNSPQKPNDRDHKRLANGLRGDVGWQVRFQFGDICYEY 362
 DB 336 LTGATHTHQYGEVCRYLHTHPISKEKMKVKAANGRLPDWQFRKRNFEYGEY 395
 QY 363 AATGNTGFMNTR--KYGAVRVNTLQKITYD--LTKDVEKDEPV-RDENGCVYR 416
 DB 396 AATEAPFATTTFOKDFGIGACRNYGTIIQWFLFOOTLVKMPDNDVSYRNSKFCFV 455
 QY 417 VPEGVGLLYCKI-----TQTPENGYAGAKAQEKKKLDVFKKGLYVNSDGLMVDHE 472
 DB 456 APVGEPEEMLRFFPKKPEPTSFQGLYGNKKEKSVVDVPRRGAWRCGDLAKADXY 515
 QY 473 NFIYFHDVGDTEPRKGENNATTEADV--TVGLVDYQVENVYGVAPDHEGRIGASIK 530
 DB 516 GLWYFLDRMGDTFRKMSSENVSTVEEDQLASKKEQYAVIVGVIPYRGRAGAVIK 575
 QY 531 MKENH-EFDCK-KLFOHIAD--VLPSYARPRIRIDQTEIT 568
 DB 576 LTDNSLDTAKTKRLNDSRLNLPYAMPLEFKVDEIKMT 617

RESULT 7
 CAIC_ECOLI STANDARD; PRT; 522 AA.
 ID CAIC_ECOLI STANDARD; PRT; 522 AA.
 AC P31552;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PROBABLE COTONOBETAINE/CARNITINE-COA LIGASE (EC 6.3.2.-).
 GN CAIC.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_Taxid=562;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=O4 K74;
 RX MEDLINE=95115548; PubMed=7815937;
 RA Eichler K., Bourgis F., Buchet A., Klieber H.-P.,
 RA Mandrand-Berthelot M.-A.;
 RT "Molecular characterization of the cat operon necessary for carnitine
 RT metabolism in Escherichia coli.";
 RL Mol. Microbiol. 13:775-786(1994).
 RN [2]

OY		196	DKDVDESTEPIDESPRESVESTESTPALITYTSGTTGLPKAMATTHORIWY-GTGTLFVSGL	254
		:	: : :	:
Db		160	OLKNQ---QPALCLCAPPLSDDDPAELLFTSSTRPCKGVITHNNLRFAYSMMCAL	216
OY		255	KADVVITYITLPFYHSALLIGIHCCIVAGATTALRTFKFSASQFMDCRKRVNYIQTIGE	314
		:	: :	: :
Db		217	RDDVDYLFWMPAFHIDCOCFTAMAFAFGAFTVLVEKYSAARAFWQVOVKRAYATECPM	276
OY		315	LMLYLNSSQRKNDPDHVFR-TALGNLRDQVRKOFKRFPGDICYEFLAAETEGNIGPM-	37Z
Db		277	MIRTLWPDPSPANDQHRLREMFYLMLSROEKDAFCERFG-VRLITSYGTEITTVGIIG	335
OY		373	--NVARKGAVGRNRYLOKKITTYDLIKDYDKDFEVPDENGYCVRRPKGEVLLYCK-	428
		:	: :	: :
Db		336	DRDGKRMRPSISRGVFC-----YEAE----IRDHNK--PLPAGIGICLK	378
OY		429	ITDTLFPNNGYAACAKKOTEKKLRDFEFKGDLINRSGDLLMDHENFIYFHNRVGDIPTWK	488
Db		379	IPEKITFEKEYF-LNPQATAKL----EADGMHLTDGTGRDEEPFFYVPRRCNMILRG	432
OY		489	GENVATTEADVADLVDFVOEVENYGVHPVDEHGRIQMASIKMKEHFEDGKLFLOHAD	548
Db		433	GENVSCELEELNTIAAPKIOTDIVVGIKDISRBEAI-KAFVYLNEGETLSSEEFFRCOEJ	491
OY		549	YLPSTARPREFLRIO	562
Db		:	: : :	:
		:	: :	: :
Db		492	NMAKFVPSYLEIR	505
<hr/>				
RESULT	8			
LCFA_ECOLI	ID	LCFA_ECOLI	STANDARD:	PRT; 561 AA.
AC	P29212;			
DT	01-DEC-1992 (Rel. 24, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	LONG-CHAIN-FATTY-ACID--COA LIGASE [EC 6.2.1.3] (LONG-CHAIN ACYL-COA SYNTHETASE).			
DN	FADD OR OLDD.			
OS	Bacteriota, Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Escherichia coli.			
OX	NCBI_TaxID=562;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN-K12;			
RC	MEDLINE=94150456; PubMed=8107670;			
RA	Fulda M., Heinz E., Wolter F.P.;			
RT	"The fad gene of Escherichia coli K12 is located close to rnd at			
RT	39.6 min of the chromosomal map and is a new member of the			
RT	Amp-binding protein family.";			
RL	Mol. Gen. genet. 242:241-249(1994).			
RN	[2]			
RN	SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.			
RP	STRAIN-K12;			
RC	MEDLINE=93094273; PubMed=1460045;			
RA	Black P.N., Dirusso C.C., Metzger A.K., Helmert T.L.;			
RT	"Cloning, sequencing, and expression of the fad gene of Escherichia			
RT	coli encoding acyl coenzyme A synthetase.";			
RL	J. Biol. Chem. 267:25513-25520(1992).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RP	STRAIN-K12 / MG1655;			
RX	MEDLINE=97426617; PubMed=9278503;			
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perrin N.T., Burland V.,			
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,			
RA	Gregor J., Davis N.W., Kirpatrick H.A., Goeden M.A., Rose D.J.,			
RA	Mau B., Shao Y.;			
RT	"The complete nucleotide sequence of Escherichia coli K-12.";			
RL	Science 277:1453-1474(1997).			
RP	[4]			
RP	SEQUENCE FROM N.A.			

```

Db      253 VTFALPLTHFALEI---NCLFLFELGQNLITPNRPIGLVKELAKYPPATLGVNTL 309
Oy      316 LRYLCNSQRKRNDRDHVRLALONGL--RGDVWRQFYKRFEDICLYEYAATE----- 366
Db      310 FMAILNNKEFOQDDESSLIHLSAGSGMPVOOVAERWVKLTGYL-LLEEYGILTECAPLVSV 368
Oy      367 -----GNIGFMNARFKVGAVGVANVLLOKKITLYDLIKYDVEKDEPVDENGVCVRV 417
Db      369 NPYIDYHSSSIGL-----PVSTETAKLV-----DD---DDN----EV 399
Oy      418 PKGEVGLLVCKITIQQLPFPGNGYAGAKAQTEKKLRDVFYFGKDLYENSGDLLMWHENFIYF 477
Db      400 PPGPGGELCYKGPVY--MLGYWRDPDAT-----EIING--WLHTGDIAVMDEEGFLRI 450
Oy      478 HDRRGDFEFMRKGENVATTEADTVGGLVDVVOEVENYGVYHPDGHRIGMASIKKENHEF 537
Db      451 VDRKKMDILVSGENVYPNEIEDVYOMHPBGVQEVAAG--VPSGSGEAENVKITFYKKPDSL 508
Oy      538 DGKKLFQHADIYLPSTARPRELRIODTIEITGFHKRMKTIVEE 581
Db      509 TEESLWFCRCROLTGYKVPFLVEFRBELPKSNVNGKILRRREIDE 552

RESULT   9
ACL_VANPL ID AC ACL_VANPL STANDARD; PRT; 553 AA.
024540;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE 4-COMMARATE--COA LIGASE (EC 6.2.1.12) (4CL) (4-COMMAROTL-COA SYNTHASE).
GN OS
OS Vanilla planifolia (Vanilla).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Asparagales; Orchidaceae; Vanilla.
OX NCBI_TaxID=51239;
RN [1]
RP SEQUENCE FROM N.A.
RA Brodeur J., Xue Z.T.;
RT "Isolation and characterization of a cDNA from cell suspension cultures of Vanilla planifolia encoding 4-coumarate: coenzyme A ligase."
RL Plant Physiol. Biochem. 35:497-506(1997).
CC -1- CATALYTIC ACTIVITY: ATP + 4-COMMARATE + COA = AMP + DIPHOSPHATE + 4-COMMAROTL-COA.
CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT COVALENT BINDING OF AMP TO THEIR SUBSTRATE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X75542; CAA53230.1; -.
CC DR HSSP: P08659; ILCI.
CC DR Mendel: 24370; Vanpl.1179;24370.
CC DR InterPro: IPR000873; -.
CC DR Pfam: PF00501; AMP-binding: 1.
CC DR PROSITE: PS00455; AMP_BINDING: 1.
CC KW Ligase; Phenylpropanoid metabolism.
CC SQ SEQUENCE 553 AA; 60095 MW; 9A2D250BC84BA1CB CRC64;

Query Match          8.4%; Score 274.5; DB 1; Length 553;
Best Local Similarity 23.2%; Pred. No. 2; se 13;
Matches 118; Conservative 95; Mismatches 223; Indels 73; Gaps 21;
```

```

QY 76 DETLYTAQVDRRSNOVARALHDLRGDCVALLMGNEPARYVIMLGKGMACLN 135
DB 54 DELFTYADVELLSRRVSGSL-SKLGKODDTMILLPSPERFVAFSLFSGISTMAN 112
QY 136 YNIRAKSLHFCOCGAKVLVSPLOAVEREILSLKDDVSIYVRSSTGDSIDL 195
DB 113 PPTSTEVYKOKAKSNAKILITQ---GCYVDKVKYACENGKILISIDTTTADANIL 169
QY 196 DKVDEYSTEPIPSMRSEVTFSTPALYITSGTGLPKRAMITHORIMVGTGLTF----- 250
DB 170 -HSELTGADENEMPEKVEISPDGVVALPYSSGTGLPKVMILTKR-----GLVTSVAQ 222
QY 251 VSG-----LKKDDVYITLPEFYH-----SAALLIGHGIVGATLALTRKFSASQFMD 300
DB 223 VDBENPLMYHSDVLLCYVLPFHITSLNSVLLCGIR-----AGSGILLMKFEIYFTEL 278
QY 301 CRKYNTVIOYIGELRYLCSNPKRNDHKVRLAL-GNGLRGDVMROFYK-RFGDICI 358
DB 279 IOKYKVTIGPFPPIVLAIAKSTVVDNDLSSVRYVMSGAAPLGELEDAVRAKFPNAKL 338
QY 359 YEPYATBEGN-----IGFMN--YARKVAVGR-VNYLQKTIITDILKYDVEKDEPYRD 409
DB 339 GGGYGTETAGPYLAMCLAPAKEPFDIKSGAGCTVVRNMAEMKIV-----DP--- 383
QY 410 ENGCVYRKGVEYGLVCKITQITLTPNGYAGAKAQOTEKKKLDVFKKGLYFNSGDLIMV 469
DB 384 ETG--SSLRNHPGELICIGDQI--MKGLNDPEAT---ATIDKEG--WHTGDDIGYI 433
QY 470 DHENFYEDRVGDTFERMKGENVATTEVADTVGLDVEOVENVYGVHP--DHEGRIGM 526
DB 434 DDDDELFIYDRKELIKYGFQVAPALFALLTHPCISDAV---VPMKDEAGAEVYP 489
QY 527 ASIKMKNHEPFGCKLFOHIADYLPYSYAR 555
DB 490 AFVYKSNHNTDEDEIKOFTISKOVITPKR 518

RESULT 10
AC12_ARATH STANDARD; PRT; 556 AA.
ID 4CL2_ARATH
AC 09S725; 09J035.
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE 4-COMMATE--COA LIGASE 2 (EC 6.2.1.12) (4CL 2) (4-COMMAROYL-COA
DE SYNTHASE 2).
GN 4CL2 OR AT3G21240 OR MX18.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Magnoliophyta; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=99348176; PubMed=10417722;
RA Ehling J., Buettner D., Wang Q., Douglas C.J., Somsich I.E.,
RA Kombrink E.;
RT "Three 4-commate:coenzyme A ligases in Arabidopsis thaliana
RT represent two evolutionarily divergent classes in angiosperms.",
RL Plant J. 19:9-20(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20277480; PubMed=10819329;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones.",
RL DNA Res. 7:131-135(2000).
CC -1- CATALYTIC ACTIVITY: ATP + 4-COMMATE + COA = AMP + DIPHOSPHATE +
CC 4-COMMAROYL-COA.

```

```

CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT
CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL; AF106085; AAD47192.1; -
DR EMBL; AF106086; AAD47193.1; -
DR EMBL; AB023045; BAB01716.1; -
DR HSSP; P08659; ILC1.
DR InterPro; IPR000873; -
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP-BINDING; 1.
KW Ligase; Phenylpropanoid metabolism; Multigene family.
FT CONFLICT 247 247 W -> L (IN REF. 2).
FT CONFLICT 265 265 I -> V (IN REF. 2).
SQ SEQUENCE 556 AA; 60929 MW; E3C2ADE7529DED CRC64;

Query Match 8.3%; Score 272.5; DB 1; Length 556;
Best Local Similarity 23.3%; Pred. No. 4,2e-13;
Matches 122; Conservative 91; Mismatches 224; Indels 87; Gaps 19;

QY 77 ETLTYAQVDRRSNOVARALHDLRGDCVALLMGNEPARYVIMLGKGMACLN 136
DB 62 EYTYADVHTVSRKLAGLH-NLGVKHQDVMLLPSPFVLTFLAASFGAITSANP 120
QY 137 NIRAKSLHFCOCGAKVLVSPLOAVEREILSLKDDVSIYVRSSTGDSIDL 192
DB 121 FFPFAETSKOKAKSNAKILITV---OSRYVDKTKNLNDGVLIV---TDSALTEKCL 172
QY 193 --SFLDVEYSTEPIPSMRSEVTFSTPALYITSGTGLPKRAMITHORIMVGTGLTF 250
DB 173 RFSLELTQSEPRVDSIEKISPEDVVALP---FSSGTGLPKGMVILTKR-----GLVT 222
QY 251 -----VSG-----LKKDDVYITLPEFYH-----SAALLIGHGIVGATLALTRKFSASQFMD 299
DB 223 SVAQOVGGENENLTFNDDVILCVPMFHTYALNSIMLSIRGATILIMPKEITLLE 282
QY 300 DCRKYNTVIOYIGELRYLCSNPKRNDHKVRLA-----LGNGLRGDVMROFYKRF 353
DB 283 QIQCKKYTVAMVPPVLAIAKSPETEKYDLSYRWKSGAAPLGELEDAI---SAKF 338
QY 354 GDICIEFYATBEGN-----NIGFMN--YARKVAVGR-VNYLQKTIITDILKYDVEKD 404
DB 339 PNAKLGCGYGTETAGPYLAMSLGFAKPEPFVKSAGCTVVRNMAEMKIL-----D 387
QY 405 EPVADENGCVYRKGVEYGLVCKITQITLTPNGYAG--AKAOTEKKKLDVFKKGDYF 461
DB 388 PDTGDS-----LPRNKPGEICJINGNOI--MKGLNDPLATASTIK-----DGMV 430
QY 462 NSGDLIMVDEHNFYEDRVGDTFERMKGENVATTEVADTVGLDVEOVENVYGVHPDE 521
DB 431 HTGDVGFDDDELFIYDRKELIKYGFQVAPALFALLTHPCISDAV---VPMKDEAGAEVYP 489
QY 522 GRIGMASIKMKNHEPFGCKLFOHIADYLPYSYARPRELRIQDTI 565
DB 490 GEVVAFAVRSKDSNISSEDEIKQFVSQVYFYKRNKINVFPTDSI 533

RESULT 11
AC12_ORYSA STANDARD; PRT; 563 AA.
ID 4CL1_ORYSA
AC P17814;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)

```


Query Match	Similarity	8.3%	Score	270.5	DB 1:	Length	563;
Best Local	Similarity	24.1%	Pred.	No. 6.1e-13;			
Matches	Conservative	82;	Mismatches	234;	Indels	77;	Gaps
							21

Qy	79	LTAAVDNRKSNQVARRAL-HDHILGKRGDCVALALMGNEPAAVWMLWMLGLVYAGCAMA	CLNTYN	137
Db	65	LTAAVDNRKSRRLAALRRAPGLRGCVNMELNRSPFVLSFPAASKVGAAYTTANM		124
Qy	138	IRAKLLHLCFOCCGAKVLVSPELQAAVEIILPSLKKDVSITYVSRYSNTDIDSLDK		197
Db	125	STPHIESOLAAAGATVTVT---ESMAADKLPSSHGALGVLLID--ERRDGLHFWD		178
Qy	198	V---DEVSTEPRIPEBMSRVSFTSTALYI--YTSGTGLPKAMITHQIWIWGTGLTFVS		252
Db	179	LMSEDEAS--PLAGDEDEKVFDPDDVVALPYSSTGTGPKVMITHRSL-----STVA		231
Qy	253	-----GLKADVDVYITLPEFYH-----SALLIGIHGCI VAGATLARTKFSASQW		298
Db	232	QOVDSEENIGIHHADVILICALPMHIYSLNTIMMGCLR---VGAIVMRERFDLAAM		287
Qy	299	DDCRKYNAVTVIOIGELLRYLCLNSPQKPNDRDHKRYALGNG--LRGDVWRQFVRFQGI		356
Db	288	DLVEHRYVTIAPLRYPIVAAVAKSBAARADSLSTRVWYLSGAPNGKDIEDAFMAKLPGA		347
Qy	357	CIYEYATTEGN-----IGFMNTARKV--GAVGR-VNQLQKIIITYDLIKYDVERDEPV		407
Db	348	VLGQGYGTEGAPVLSMCIAFAKEPKFKSGAGCTVYVRAELKIIDDPDGK---SLGRNL		404
Qy	408	RDENGYCYRVAPKGEVGLLVCKITQTLTPPNGVAGAAQTEKKLRIVYFKKGDITFNSGDL		467

```

Db      405 REE--ICIR-----GQOIMK-----GYLNPNPEARK----NTIDAEG--WLHTGDIG 442
QY      468 MDHENEFLYFHRVODTFRMKGENATTEVADYGLVDFVEVAVNYGVHPDHEGRGMA 527
Db      443 YVDDDELFIVDLRELKILYKRFQVAPAEALALLNTHPSIDAAVGLKF---GEIPVA 498
QY      528 SIKMEKNEHFDEGKLFQHIADILPESVAPRFLRIODTI 565
Db      499 FVAKTEGSELSDDYKQFVAKEVIIYKKIRREVEFYDKI 536

RESULT 12
ID      SRFL_BACSU      STANDARD;      PRY;      3587 AA.
AC      P27206;
DT      01-AUG-1992 (Rel. 23, Created)
DT      01-FEB-1995 (Rel. 31, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      SURFACTIN SYNTHETASE SUBUNIT 1.
GN      SREFA OR SREFAI OR SREFA.
OS      Bacillus subtilis.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group;
OC      Bacillus/Staphylococcus group; Bacillus.
OX      NCBI_TaxID=1423;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=168;
RX      MEDLINE=93181186; PubMed=8441623;
RA      Fuma S., Fujishima Y., Corbell N., D'Souza C., Nakano M.M.,
RA      Zuber P., Yamane K.:
RT      "Nucleotide sequence of 5' portion of srfA that contains the region
RL      required for competence establishment in Bacillus subtilis.";
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=168 / JH642;
RX      MEDLINE=93360813; PubMed=8355609;
RA      Cosmana P., Rodriguez F., de Ferra F., Grandi G., Perrego M.,
RA      Yenema G., van Sinderen D.:
RT      "Sequence and analysis of the genetic locus responsible for surfactin
RL      synthesis in Bacillus subtilis.";
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=168;
RX      MEDLINE=97124189; PubMed=8969502;
RA      Yamane K., Kumano M., Kurita K.:
RT      "The 25 degrees-36 degrees region of the Bacillus subtilis
RL      chromosome: determination of the sequence of a 146 kb segment and
RN      identification of 113 genes.";
RN      [4]
RP      SEQUENCE OF 1-460 FROM N.A.
RX      MEDLINE=91154134; PubMed=1847909;
RA      Nakano M.M., Magnusson R., Myers A.M., Curry J., Grossman A.D.,
RA      Zuber P.:
RT      "srfA is an operon required for surfactin production, competence
RL      development, and efficient sporulation in Bacillus subtilis.";
RN      [5]
RP      SEQUENCE OF 1-38 FROM N.A.
RX      MEDLINE=91358326; PubMed=1715856;
RA      Nakano M.M., Xia L., Zuber P.:
RT      "Transcription initiation region of the srfA operon, which is
RL      controlled by the comp-comA signal transduction system in Bacillus
RN      J. Bacteriol. 173:5487-5493(1991).
RN      [6]
RP      SEQUENCE OF 1-64 FROM N.A.
RC      STRAIN=168;
RX      MEDLINE=95219080; PubMed=7704255;
RA      Fujishima Y., Yamane K.:
RT      "A 10 kb nucleotide sequence at the 5' flanking region (32 degrees)

```

[illegible]

01-APR-1990 (Rel. 14, Created)
 01-APR-1990 (Rel. 14, Last sequence update)
 01-OCT-2000 (Rel. 40, Last annotation update)
 4-COUMARATE--COA LIGASE 1 (EC 6.2.1.12) (4CL 1) (4-COUMAROYL-COA
 SYNTHASE 1).
 GN 4CL1 OR 4CL-1.
 OS Petroselinum crispum (Parsley) (Petroselinum hortense).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicots; Asteridae;
 CC eusterids II; Apiales; Apiaceae; Petroselinum.
 OK NCBI_TaxID:4043;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE:89005119; PubMed:3169018;
 RA Locoya E., Hoffmann H., Douglas C., Schulz W., Scheel D.,
 RA Hahlbrock K.,
 RT "Primary structures and catalytic properties of isoenzymes encoded by
 RT the two 4-coumarate:COA ligase genes in parsley.",
 RL Eur. J. Biochem. 176:661-667(1988).
 RN [2]
 RP SEQUENCE OF 1-8 FROM N.A.
 RA Douglas C., Hoffmann H., Schulz W., Hahlbrock K.,
 RT "Structure and elicitor or U.V.-light-stimulated expression of two
 RT 4-coumarate:COA ligase genes in parsley.",
 RL EMBO J. 6:1189-1195(1987).
 CC -1- CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA = AMP + DIPHOSPHATE +
 CC 4-COUMAROYL-COA.
 CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
 CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
 CC -1- INDUCTION: TRANSCRIPTIONALLY ACTIVATED BY BOTH FUNGAL ELICITOR AND
 CC ULTRAVIOLET IRRADIATION.
 CC -1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT
 CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X13324; CAJ31696.1; -;
 CC EMBL: X05350; CAJ28959.1; -;
 CC PIR: S01667; S01667.
 CC HSSP: P08659; ILCI.
 CC InterPro: IPR000873; -;
 CC Pfam: PF00501; AMP-binding: 1.
 CC PROSITE: PS00455; AMP-BINDING: 1.
 CC KW Ligase; Phenylpropanoid metabolism; Multigene family.
 CC SQ SEQUENCE 544 AA; 59825 MW; 22BAD78F255D0C8 CRC64;

 Query Match 8.18; Score 265.5; DB 1; Length 544;
 Best Local Similarity 24.48; Pred. No. 1,4e-12;
 Matches 111; Conservative 75; Mismatches 183; Indels 85; Gaps 20;

 QY 77 ETLTYAODRKSNOVARALHDHGLRQDCVALLMGNEPVMVLGLVLCGAMCLNY 136
 DB 52 ETLTYOVELLSRKVAVSGL-NKLGIOGDDIMLLRPSPEYFAFLGASIRGALSTMANP 110
 QY 137 NIKRSLHCFQCCGAVLVLPDLAAVEELPSLKKDVSIIYYSRTSNTDGDLS-- 193
 DB 111 FETSAEVIKOLIKASQAKLI-----ITQACYVDKDYAAAKNKIICIDDAQ 158
 QY 194 ---FLDLYDEVSSTPIPE-SMRSEVTSPTALVYTSSTGLSKRANITQRWYSGGLT 249
 DB 159 DCLHFSKLMRDESEMEPEVYNSDDVALP---YSSGTTGLPRGVALLTKR-----GLV 208
 QY 250 F-----VSG-----LKADVITYITLPEYH-----SAALLIGIGCIYAGATLALRFFSA 294
 DB 209 TSVAGQVDGDNPNLYHMSDEVMCIILPLFIHYSLNVLNVLGGLR-----AGVTILIMKRFDI 264

 QY 295 SQFWDCRKYNVTVIOTIGELRLYCNSPKPNDRHKYRLAL-GNGLRGDWVRQFVK-R 352
 DB 265 VPELELLQKRYKVTIGPEVPIVLAIAKSPYVDKYLSSVFTWMSGAPLGELEDAVRAK 324
 QY 353 FGDCIYEFYAATGEGN-----IGFMN--YARVGAAGR-VNLTQKKITTYDLIKDVAK 403
 DB 325 FPNNAKGGGIGMTBAGPYLAMCIAFAKEPEIKSGAGGVVRAEMKIV-----DPEP 377
 QY 404 DEPVRENGCYCVRPGEVGLVCKITQTLTPFNGVAGAAQTEKKLRVFKKDLDFNS 463
 DB 378 N-----ASLPRNGREICIRGDDI--MKGIINDESTR-----TIIDEEG--WLT 419
 QY 464 GDLLMDHENTFITFDVRGDTFPRKGENVATTEV 497
 DB 420 GDIGFIDDDDELPIVDRLKEIIRKGFQVAPAEI 453

 RESULT 14
 ID LCFA-HAEN STANDARD; PRT; 562 AA.
 AC P46450;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE LONG-CHAIN-FATTY-ACID-COA LIGASE (EC 6.2.1.3) (LONG-CHAIN ACYL-COA
 SYNTHETASE).
 GN FADD OR HD390.1.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Haemophilus.
 OK NCBI_TaxID:727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RD / KW20 / ATCC 51907;
 RX MEDLINE:95350630; PubMed:7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-T., Glodak A., Kelley J.M.,
 RA Weidman J.E., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utechtack T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus
 RT influenzae Rd.",
 RL Science 269:496-512(1995).
 RN [2]
 RP IDENTIFICATION.
 RA Koonin E.V., Rudd K.E.;
 RL Submitted (SEP-1995) to the SWISS-PROT data bank.
 CC -1- FUNCTION: ESTERIFICATION, CONCOMITANT WITH TRANSPORT OF EXOGENOUS
 CC LONG-CHAIN FATTY ACIDS INTO METABOLICALLY ACTIVE COA THIOESTERS
 CC FOR SUBSEQUENT DEGRADATION OR INCORPORATION INTO PHOSPHOLIPIDS
 CC (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + A LONG-CHAIN CARBOXYLIC ACID + COA = AMP
 CC + PYROPHOSPHATE + AN ACYL-COA.
 CC -1- COFACTOR: REQUIRES MAGNESIUM (BY SIMILARITY).
 CC -1- SUBUNIT: HOMODIMER (PROBABLY).
 CC -1- SUBCELLULAR LOCATION: PARTIALLY MEMBRANE-ASSOCIATED (POTENTIAL).
 CC -1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT
 CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U32722; AAC22049.1; -;
 CC HSSP: P08659; ILCI.

DR TIGR:HI0390.1; -.
DR InterPro: IPR000673; -.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00455; AMP-BINDING; 1.
KW ligase; Fatty acid metabolism; Magnesium; Membrane.
SQ SSEQUENCE 562 AA; 63478 MW; 74BEFC89D711P12 CRO64;

Query Match	8.18;	Score 264.5;	DB 1;	Length 562;
Best Local Similarity	23.48;	Pred. No. 1.7e-12;		
Matches 14;	Conservative	87;	Mismatches 265;	Indels 87;
				Gaps 20;

QY	54	TIRATLEAAROTPKPFLTFREDETLTYAODVRSNOVARALDHJGROCVALLMGN	113
Db	24	STLMDKAVREHPRPRAYIMKGOVLTFRKLEERSRAFAAYLONEKTKLQROGRALAMPN	83
QY	114	EPAYVWMLGVLKVCACMACLNMYNIRAKSLHCFQCCAKVLVSPLOAAVEETLPSLK	173
Db	84	LLOYPALGLIIRAGLIVANNPLTFRELELOLOSAGAVLVVSNFASLTLEKV----	139
QY	174	KDOVSITYVSPISNMDGD-----SFLDKV-----DEVSFEPIPS	209
Db	140	FNTNKHVITLRMGDQLSFGKRLVNFVVKYKKVLPRYKTLPHAVTFEVLSTGKTYQY	196
QY	210	WRSEVEFSPALYIYTSGTGLPRAMITTHQRI-----WYTGELTFVSGLRADVY	261
Db	199	VREISREDDLAFLQYTGCTGVAKGAMLTGHNITITVFOAKWIAE--PRIGHSRTSAAI	256
QY	262	ITLPRFSAALLIGHGCVAGATLALRTK-FASQOWDQCKRYNTVYQYIGELRLTC	320
Db	257	LAPRLHVALTIVNCLLELGLVATLITLTPRDIEGVEKLKRYREALTYNTLNAL	316
QY	321	NSFOKPNDRDHVKVRLALNGSLRGDWAPVYKREGDI--C-IYEFYAATEGNIGF----	372
Db	317	NNEMFEKVPFSALKLSVGGMA--IQGSVATRMHLELGCNIIEGYMTCECSPLIACPIN	374
QY	374	YARKVGAV--RNVYLQKKIITTYDLTKIDVEDEYVDENGYCYRVKGVGLVYCKITQL	433
Db	375	VVHNSTGTIGAPPNPNDIKI-----KDD-----GSDAKI--GEAGELWKGQV	416
QY	433	TPRNGAGAAKAOETEKKLLIDVFKKGDLYFNSGDILMWDEHNFYFHDYRGDFRWMGENV	492
Db	417	--WRGWQRPENTS-----EVLAKDS--WMATGDIWIDESYLRIDRKKDILLYSGENV	466
QY	493	ATTEVADPVGVLVDQEVANNYGV--HVPDEHGRIGMASIK--MKNHEFDGKILFQITAD	546
Db	468	YPRIEIDVWLVNMYKSEAVALGVPAAVSGE-----TIKIFVYKKDDSLTRDELPHNRQ	521
QY	549	YLPSPARPRRLRIODTIETGTFFKRRKMTLVEE	581
Db	522	YLTGYVPEKIEPRDELPRXTNVGKILRRVLKDE	554
RESULT 15			
ACL_PINTA	15		
ID	ACL_PINTA	STANDARD:	PRT: 537 AA.
AC	PA1636:		
DT	01-NOV-1995 (rel. 32, Created)		
DT	01-NOV-1995 (rel. 32, Last sequence update)		
DT	01-OCT-2000 (rel. 40, Last annotation update)		
DE	4-COMPARTE--COA LIGASE (EC 6.2.1.12) (4CL) (4-CODUAROYL-COA		
DE	SYNTHASE).		
GN	4CL.		
OS	Pinus taeda (loblolly pine).		
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;		
OC	Coniferopsida; Coniferales; Pinaceae; Pinus.		
OX	NCBI_TaxID=3352;		
RN	1		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=L.; TISSUE=Xylem.		
RA	voe K.S.;		
RL	submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.		
LN	[2]		

RP SEQUENCE FROM N.A.
 RC TISSUE=Needle;
 RX MEDLINE=97161117; PubMed=9008388;
 RA Zhang X.H., Chiang V.L.;
 RT "Molecular cloning of 4-coumarate:coenzyme A ligase in loblolly pine
 RT and the roles of this enzyme in the biosynthesis of lignin in
 RT compression wood.";
 RL Plant Physiol. 113:65-74(1997).
 CC -1- CATALYTIC ACTIVITY: ATP + 4-COUMARATE + COA = AMP + DIPHOSPHATE +
 CC 4-CODMAPROYL-COA.
 CC
 CC -1- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
 CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
 CC -1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT
 CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.
 CC
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

[illegible]

Db 454 ALLVAHPSTADAAV---VPQKHEEAGEVPVAFVYKSSSEISEQETKEFVAKOVIFKKIH 509
QY 544 -QHADYLPSTYARPRFLR 560
Db 510 RYFVDAIPKSPSGKILR 527

Search completed: July 16, 2001, 18:15:11
Job time: 187 sec

2